

LASER SCANNING AND AUTOMATIC MULTI-IMAGE TEXTURING OF SURFACE PROJECTIONS

E. Petsa¹, L. Grammatikopoulos^{1,2}, I. Kalisperakis^{1,2}, G. Karras², V. Pagounis¹

¹Department of Surveying, Technological Educational Institute of Athens (TEI-A), GR-12210 Athens, Greece

²Department of Surveying, National Technical University of Athens (NTUA), GR-15780 Athens, Greece

E-mail: petsa@teiath.gr, laazaros@central.ntua.gr, ilias_k@central.ntua.gr, gkarras@central.ntua.gr, mail@pagounis.gr

KEY WORDS: Automation, Heritage Conservation, Laser Scanning, Orthorectification, Texture

ABSTRACT

Terrestrial laser scanning has now become a standard tool for 3D surface modeling. In order to exploit such fully 3D data in texture-mapping or for the creation of large-scale ‘true orthos’ suitable software is needed, particularly to allow handling surface self-occlusions and image occlusions, as well as multi-image texture interpolation. The authors have presented such an automatic approach for creating orthoimages and perspective views, based on fully 3D models from laser scanning. All surface points visible in the direction of projection are first identified, and then texture from all images which view each particular surface point is blended. Means for automatically eliminating colour outliers from individual images, especially near image occlusion borders, are also provided. In this contribution, the algorithm is evaluated using image sets with large variations in image scale and unconventional imaging configurations. The presented results indicate that this approach, which involves a cooperation of photogrammetry with laser scanning for the automatic multi-view synthesis of textured projections, performs quite satisfactorily also under demanding circumstances.

1. INTRODUCTION

In the last years, the synergy of photogrammetry with terrestrial laser scanning (TLS) has become a reality which attracts intense interest in the field of cultural heritage, in particular regarding photo-texturing of 3D surface models (e.g. Aguilera *et al.*, 2006; Alshawabkeh & Haala, 2005; Zhang *et al.*, 2006). Admittedly, TLS is not a panacea – depending on object morphology or cost issues etc. – and hence it cannot totally replace image-based 3D modeling. Yet in numerous instances it allows fast and accurate surface modeling, i.e. a fully 3D basis for texture mapping, and for orthoprojection in particular.

This availability of fully 3D models has challenged conventional orthorectification tools, especially regarding issues of occlusion. Opposite to the use of a 2.5D model, surface self-occlusion in the direction of projection becomes now important; besides, image occlusion must also be recognized. In this sense, rigorous digital orthoprojection tools must handle both aspects of occlusion. Although the concern of this paper is orthoprojection, it is emphasized that the occlusion/visibility issue refers to the much broader context of 3D texture-mapping (e.g. Früh *et al.*, 2004).

In addition, single-image texturing often suffers from abrupt radiometric changes or ‘discontinuity artefacts’, i.e. visual degradations of the new image. Consequently, it is advisable to blend colour from more images in a weighted average of corresponding values (Baumberg, 2002; Orzan & Hasenfratz, 2005). In the case of photogrammetric mapping it seems reasonable to create one single texture map, with each source image contributing according to its spatial relation to the object (Poulin *et al.*, 1998). Tools for excluding wrong colour values are also needed.

The authors have presented a multi-view algorithm for automatically creating orthographic or perspective views based on existing 3D models (Grammatikopoulos *et al.*, 2005). Elaborations on the approach – mainly concerning the occlusion issue – were reported recently (Grammatikopoulos *et al.*, 2007; Karras *et al.*, 2007). The main intention of this contribution is to evaluate the performance of this automatic algorithm in demanding practical circumstances (unconventional imaging configurations, objects with large extension in 3D, automatic creation of cross-sections) and to discuss in more detail issues of practical significance.

2. OUTLINE OF THE ALGORITHM

2.1 Visibility checking and texture blending

Briefly, our algorithm establishes occlusions in two steps, as reported in Grammatikopoulos *et al.* (2005):

- *Surface parts visible in the given (orthogonal) projection.* The 3D mesh is projected onto the projection plane. Identification of all projected model triangles containing an orthoimage pixel means that all 3D triangles intersected by this projection ray are known. Among all intersection points, that closest to the projection plane should be visible on the orthoimage. Thus, an orthoimage depth map is generated (all pixels of the orthoimage are uniquely associated with one corresponding model point).

- *Areas of each source image entitled to contribute texture to the model parts visible in the projection.* Model triangles are projected centrally onto all source images. For each orthoimage pixel its corresponding model point is projected onto all images. The 2D triangles containing the resulting image points are identified; hence, all 3D triangles met by the projection ray of each image may be established. The intersection point nearest to the respective projection centre is that actually seen on each image. As a result, it is known which images may contribute colour for a particular orthoimage pixel and which must not.

Each image may now be resampled to create an auxiliary ‘true’ orthoimage, with occluded areas left blank (matters referring to colour interpolation in 3D are discussed in Karras *et al.*, 2007). Colour value for each pixel of the final orthoimage is the mean of corresponding values from the non-blank areas of all images, weighted with the surface area of the corresponding 2D triangle (Poulin *et al.*, 1998). This weighting scheme takes into account the viewing angle and image resolution in model space. Yet, if images differ largely in scale, the unconditional participation of all viewing images in blending may eventually produce blurred final textures. This might be avoided e.g. by using the squared surface area of 2D triangles or by blending colour only from the N images with highest weight (Grammatikopoulos *et al.*, 2007).

2.2 Removal of texture ‘blunders’

Colour values from different images to be blended for texturing an orthoimage pixel may well include outliers, stemming from

obstacles, or orientation and modeling errors which may have strong effects near occlusion borders. In order to eliminate them before assigning texture to some orthoimage pixel, Bornik *et al.* (2001) apply a median filter for detecting values outside a user-specified range. By using the mean μ and standard deviation σ from all valid colour values for each orthoimage pixel in order to discard those outside the range $\mu \pm \sigma$ (σ may also refer to the median which is more sensitive to outliers), Grammatikopoulos *et al.* (2005) have presented successful removal of occlusion-induced artefacts and obstacles. However, it seems that such basic tests are not always effective in cases of scattered colour values. Furthermore, these tests necessitate at least three colour values.

If for a pixel of the orthoimage texture from only two images is available, Karras *et al.* (2007) suggested to exploit image depth maps for evaluating the 'occlusion risk' of a pixel, based on a comparison of its imaging distance with those of its neighbours. Alternatively, imaging distances can be used to introduce image 'weight maps' (Bamberg, 2002). Yet, changes in depth maps do not necessarily imply existence of occlusion borders.

Thus, Grammatikopoulos *et al.* (2007) have proposed a straightforward single-image treatment of occlusion borders relying on morphological transformations and edge extraction. All source images are resampled to generate corresponding ('true') orthoimages, in which occluded regions are represented by a specific colour; these are then binarized to single out occluded areas. An opening operator is applied to the binary image to remove small areas of a few pixels. Image occlusion edges on this orthoimage are extracted via a Laplace filter. Subsequently these 3D entities (each orthoimage pixel has a depth value) are back-projected on the original image and expanded with a dilation operator. In this mode, occluded areas of the original image have slightly grown. The range of the dilation step depends on the assumed precision of 3D modeling and image orientation. The contribution of each image in the final orthoimage will not include texture from the regions added by this expansion of occlusion edges. Of course, the final texture blending process will still only involve colour values which pass the $\mu \pm \sigma$ outlier test.

3. EXPERIMENTAL APPLICATIONS

3.1 Test object

Purpose of the applications performed here was to test the algorithm under practical conditions more demanding than those of our previous publications. Therefore, the chosen object showed large differences in depth, combined with significant occlusion problems. Besides, it had to be recorded with image sets largely differing in scale in order to ensure possibly uniform texture input, which poses certain problems regarding image registration. A part of a Roman archaeological site in Arcadia was scanned with a Mensi GS200 laser scanner in 9 scans at 1 cm resolution. To bypass some difficulties met during scan registration with the available software, the 3D similarity transformation

$$\mathbf{X}_i = \mathbf{R}_k \mathbf{x}_{ik} + \mathbf{t}_k$$

was used to transform in a one-step solution all control and 'tie' points \mathbf{x}_i of scans k in the geodetic system (\mathbf{X}_i) through rotations \mathbf{R}_k and translations \mathbf{t}_k of each scan system (the overall precision of registration was ± 7 mm).

Tree parts (1, 2, 3) of the 3D model, seen in Fig. 1, were chosen. The sub-models consisted of 3.5, 3.6 and 0.8 million triangles, respectively. The first two parts had large differences in depth, the third was relatively shallow. Part 1 includes a semi-cylindrically shaped surface, which represents a challenge for orthoprojection algorithms. In order to ensure that different object areas

would not be textured with very unequal input resolutions, two large image sets with varying imaging distances and orientation were used. The first was taken with a 7.2 Megapixel camera and consisted of 16 images (Fig. 2), the second included 23 images taken with a 10 Megapixel camera (Fig. 3).

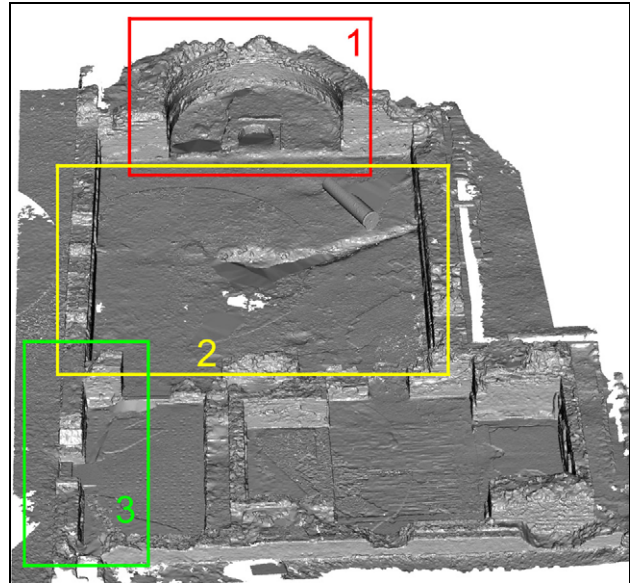


Figure 1. Full 3D model and sub-models 1, 2 and 3.



Figure 2. First image set of part 1.

Part 2 (which included two occluding planes) was recorded on 17 images with the 7.2 Megapixel camera, acquired from much

differing imaging distances (Fig. 4). Finally, Part 3 was taken with the same camera on 10 images, also with large differences in imaging distances (Fig. 5).



Figure 3. Second image set of part 1.

3.2 Bundle adjustment

In texture-mapping, the good agreement among textures coming from different images is clearly an important requirement. This is even more the case when large-scale projections products are to be generated by blending colours from multiple views. Thus, it is strongly advisable to secure that all images are tightly registered to each other as well as to the 3D surface model, which implies that all images to be used for texturing should participate in a single self-calibrating bundle adjustment. Clearly, image configurations which are suitable for securing good textures for the whole object do not necessarily represent an ideal geometry for bundle adjustments. Such configurations may be weak – e.g. due to lack of convergence between image axes, due to similar image strips taken from different distances simply for ensuring similar resolutions in different depths (which gives acute angles of ray intersection), due to small variations in 3D. In these cases it is unavoidable that estimated parameter values will be of low precision (highly correlated). Such estimates may not be at all reliable, e.g. as a camera calibration data, yet on the other hand they represent an optimal agreement among all available data.

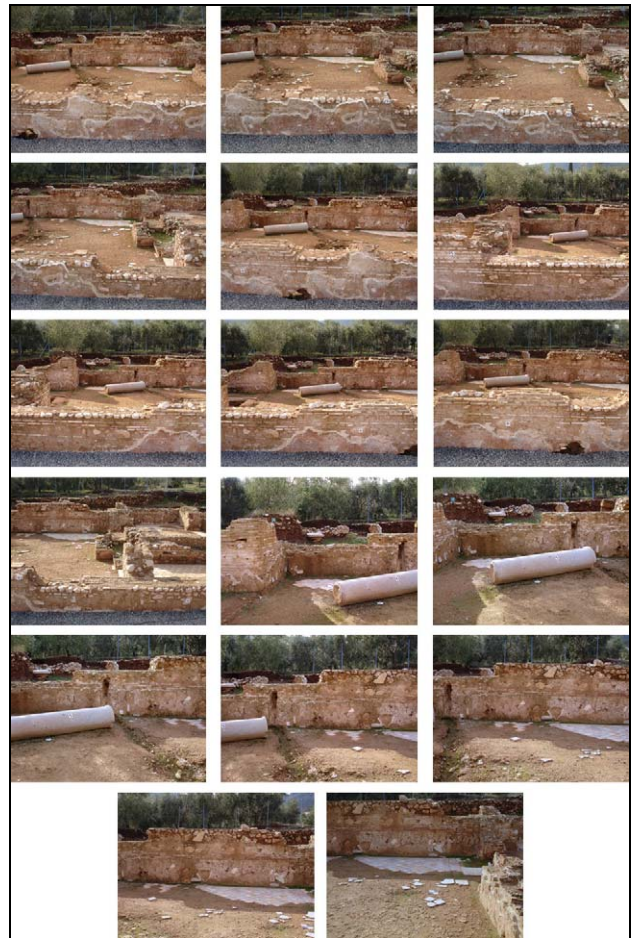


Figure 4. Image set of part 2.



Figure 5. Image set of part 3.

This may be illustrated at the example of the three image sets from the same camera. Results for their adjustments are seen in Table 1. Estimates for the camera principal point location differ considerably among adjustments and are quite uncertain. This is particularly true for part 3 in which object relief is small. In this last case the camera constant is estimated with an uncertainty of 5%. Yet, although these results do not reflect a reliable camera

calibration, this did not affect ray intersections significantly, as was seen in the texturing process.

	part 1	part 2	part 3
# of images	17	16	10
σ_o (pixel)	0.54	0.61	0.58
c (pixel)	3418.4 ± 4.9	3415.2 ± 6.3	3409.7 ± 17.4
x_o (pixel)	6.6 ± 3.1	2.6 ± 3.2	57.9 ± 10.3
y_o (pixel)	-13.5 ± 7.3	-3.7 ± 4.3	26.4 ± 13.3
k_1 ($\times 10^{-9}$)	-5.2 ± 0.2	-5.4 ± 0.3	-5.8 ± 0.6
k_2 ($\times 10^{-15}$)	-1.2 ± 0.1	-1.2 ± 0.1	-1.1 ± 0.2

On the other hand, it is to note that the estimation of radial lens distortion is quite repeatable, as seen in the calibrated curves of Fig. 6. Finally, all standard errors of the adjustments exceeded slightly 0.5 pixels, which is attributed to the fact that the control points had not been measured geodetically but were simply extracted from the point cloud. The final products of the texturing process indicated that for purposes of precise texture-mapping one may trust control points extracted directly from the 3D data.

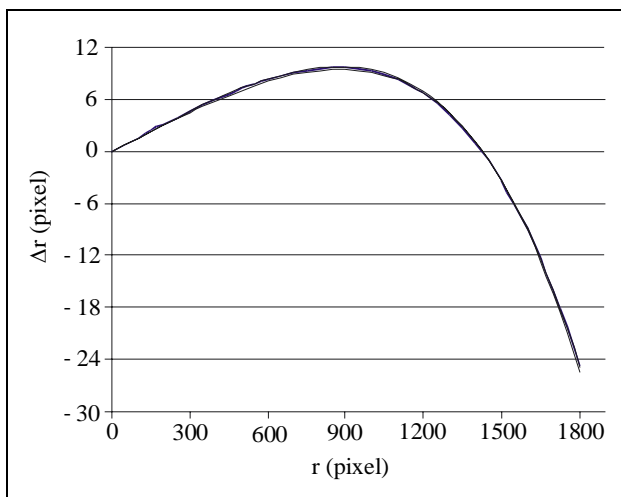


Figure 6. Calibrated curves of radial lens distortion.

3.3 Textured products

After histogram matching, all images of each block participated in the orthoimage generation. The pixel size in object space was fixed to 2 mm, a resolution suitable for the 1:20 scale. As noted, colour blending from images with different scales, i.e. different resolutions in object space, is expected to produce a somewhat blurred ('muddy') result. If only the N best images – those with largest weights – are allowed to participate, the resulting image has more detail and no radiometric discontinuity, yet it might be more 'noisy'. Fig. 7 gives such an example for $N=3$. It appears that, in each case, one should finally choose the procedure best adapted to the specific image set.

Fig. 8 shows the orthoimage from a source image and the back-projection of the occlusion borders onto the original image. The described approach for the automatic extraction and subsequent dilation of image occlusion borders has improved results. It is seen in Fig. 9, which shows an orthoimage detail of part 3, that the edge is very clearly outlined. The edge is equally clear in Fig. 10, which presents an orthoimage detail of part 1. The edge is generated here from 9 images, i.e. from all images among the 23 of the image set which view this particular region. In Fig. 11 the area of another edge, again of part 1, is shown as generated from the two different image blocks. This particular object area is orthoimaged in both cases from the 8 images which view it.

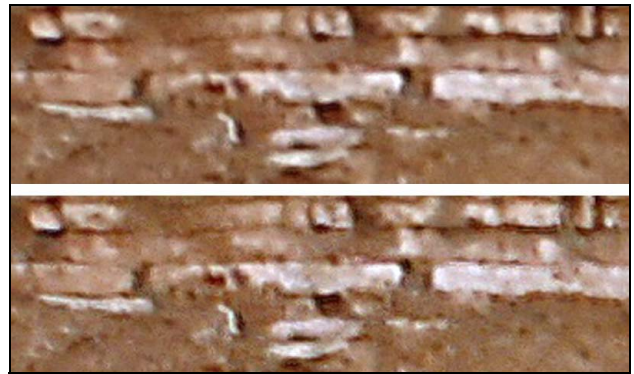


Figure 7. Orthoimage details from all images (top) and from the three 'best' images (bottom).



Figure 8. Orthoimage from one image (top) and projection of its occlusion borders onto the original image (bottom).



Figure 9. Orthoimage detail for part 3 from all images.

It is pointed out that in Fig. 11 the column in the foreground, as seen in the image sets of Figs. 2 and 3, has been 'jumped over' (see also Fig. 13). Indeed, provided that 3D data and texture are available, the algorithm can automatically create cross-sections by simple depth thresholding (Grammatikopoulos *et al.*, 2005). A further example from part 2 is seen in Fig. 12.



Figure 10. Orthoimage detail of part 1 generated by blending texture from 9 images.

Fig. 13 shows the full orthoimages from the two image sets of part 1. Deformations in the top left corner of the lower image are explained by the fact that a part of the monument had been covered for protection between the two recordings. Finally, Fig. 14 presents the orthoimage of part 2 and the projection of the posterior wall created by depth thresholding.

5. CONCLUDING REMARKS

On the assumption that accurate 3D surface mesh and image calibration/orientation data for overlapping images are given, our algorithm identifies object and image occlusion to synthesize in a fully automatic multi-image mode textured projections of the 3D model. Besides orthographic, these may also be perspective (Grammatikopoulos *et al.*, 2005) and cross-sections. Pixels are coloured through weighted participation of all viewing images.

In our experience, most cases of wrong colouring are spotted by the simple blunder detection tool, yet further means are also required, such as the morphological operators used here for minimizing the effects of occlusion borders on texture-blending.



Figure 11. Orthoimage details of part 1 generated by blending texture from 8 images of two different images sets.

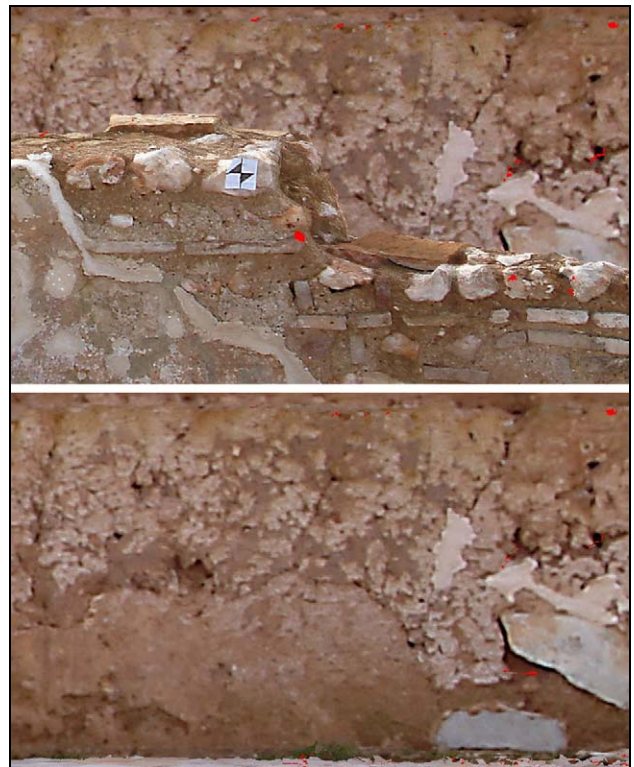


Figure 12. Detail of the oOrthoimage of part 2 (top) and corresponding cross-section (bottom).

The presented results are regarded as quite satisfactory, and the algorithm behaves very well in the area of edges. Main defects are small holes due to lack of texture. As noted in Grammatikopoulos *et al.* (2007), this problem is artificially aggravated by

the dilation operator used for expanding occlusion edges, which also increases all existing small clusters of occluded pixels. An elaboration would be necessary (e.g. by connected components labeling) for the removal of such occlusion 'blobs'. In any case, the primary role of the precision of original input data remains. Besides the accuracy of 3D modelling, this involves the reconstruction of the imaging configuration, i.e. image blocks have to be processed in a single self-calibrating adjustment.

REFERENCES

Aguilera D.G., Lahoz J.G., Finat J., Martínez J., Fernández J., San José J. (2006). Terrestrial laser scanning and low-cost aerial photogrammetry in the archaeological modeling of a Jewish tanneries. *Int. Arch. Photogrammetry, Remote Sensing & Spatial Information Sciences*, 36(5), pp. 8-13.

Alshawabkeh Y., Haala N., 2005. Automatic multi-image phototexturing of complex 3D scenes. *XVIII CIPA Int. Symposium*, Torino, 27 September - 1 October, pp. 68-73.

Baumberg A., 2002. Blending images for texturing 3D models. *British Machine Vision Conference*, pp. 404-413.

Bornik A., Karner K., Bauer J., Leberl F., Mayer H., 2001. High quality texture reconstruction from multiple views. *Journal of Visualisation & Computer Animation*, 12(5), pp. 263-276.

Früh C., Sammon R., Zakhor A., 2004. Automated texture mapping

of 3D city models with oblique aerial imagery. *2nd Int. Symposium on 3D Data Processing, Visualization and Transmission*, Thessaloniki, pp. 396-403.

Grammatikopoulos L., Kalisperakis I., Karras G., Petsa E., 2005. Data fusion from multiple sources for the production of orthographic and perspective views with automatic visibility checking. *XVIII CIPA Int. Symposium*, Torino, pp. 819-824.

Grammatikopoulos L., Kalisperakis I., Karras G., Petsa E., 2007. Automatic multi-view texture mapping of 3D surface projections. *3D Virtual Reconstruction & Visualization of Complex Architectures (3D Arch '07)*, July 12-13, ETH Zürich.

Karras G., Grammatikopoulos L., Kalisperakis I., Petsa E., 2007. Generation of orthoimages and perspective views with automatic visibility checking and texture blending. *Photogrammetric Engineering and Remote Sensing*, 73(4), pp. 403-411.

Orzan, A., Hasenfratz, J.-M., 2005. Omnidirectional texturing of human actors from multiple view video sequences. *IEEE Int. Workshop on Human Computer Interaction*, Beijing.

Poulin P., Ouimet M., Frasson M.-C., 1998. Interactively modeling with photogrammetry. *Eurographics Workshop on Rendering '98*, pp. 93-104.

Zhang Y., Yana L., Lu X., 2006. Precise texture modeling with 3D laser scanning. *Geoinformatics 2006: Remotely Sensed Data and Information*, Proc. SPIE, vol. 6419, paper no. 64191Z.

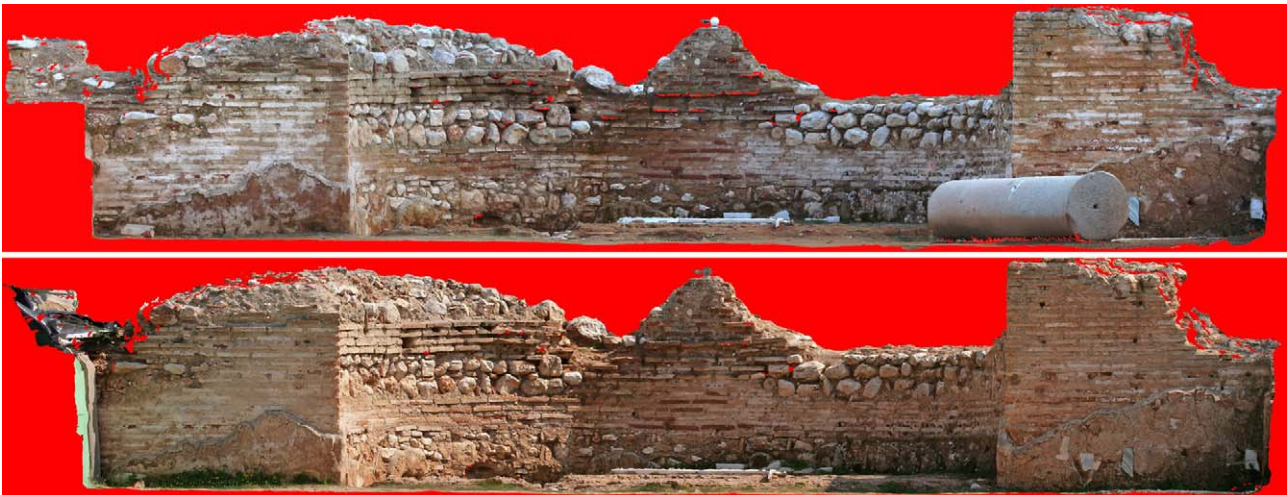


Figure 13. Orthoimages of part 1 from the two image sets (top: 16 images, bottom: 23 images).



Figure 14. Orthoimages of part 2 from 17 images (top). Below is the projection of the posterior wall from the same images.